A Conversational Multi-Agent Platform for BIM Data Intelligence

Amirhossein Nourbakhshrezaei[1], Nolan Porther[1], Mojgan Jadidi[1,*]

[1]Lassonde School of Engineering, York University, 4700 Keele Street, Toronto, ON M3J 1P3, ON, Canada,
amirnbr2@yorku.ca, nolanpo@yorku.ca, mjadidi@yorku.ca

Keywords: Building Information Modelling (BIM), Multi-Agent Systems (MAS), Conversational, Artificial Intelligence (AI), Semantic and Geometric Reasoning, Natural Language Interaction, Automated BIM Analysis

Abstract

This paper proposes the development of a multi-agent system (MAS) for Building Information Modeling (BIM) environments, where users interact with a 3D model and a chat-bot to query, validate, and analyze building elements. By leveraging conversational AI and modular agents capable of semantic understanding and geometric computation, this system allows users to retrieve data, perform quality checks, and visualize computed results directly using the BIM information. The approach supports diverse tasks, from attribute completion and filtering to volumetric calculations, thus enabling a more intelligent and accessible BIM experience for analytical purposes.

1. Introduction

Building Information Modeling (BIM) has transformed the construction and facility management industries by enabling detailed digital representations of built assets. Despite the advantages, interacting with complex BIM requires technical expertise and manual data processing. Recent advances in Artificial Intelligence (AI), particularly Large Language Models (LLMs), offer a new paradigm enabling natural language interaction with structured, geometrical, and spatial data. This paper presents a novel multi-agent system that connects users to a BIM through a chat interface. Each agent is designed to perform specific tasks, such as data retrieval, attribute validation, spatial analysis, and geometrical computation. The goal is to bridge the gap between domain experts and BIM data by allowing intuitive exploration and automated insights generation.

2. Problem Definition

BIM users often encounter several key challenges: Manual querying of attributes and geometric data is time-consuming. BIM contains extensive data on building components, materials, etc., which can be overwhelming to users. Users from various disciplines may not be familiar with specific terminology, naming conventions, and software used for BIM. Detecting missing or invalid attributes across large models is nontrivial. Volume, area, and length calculations often require specialized knowledge.

3. Background

Natural language interfaces for BIM have been explored to make building data more accessible. Early approaches used predefined ontologies and semantic parsers to map user queries to BIM data, but these were limited in scope and required rigid schemas. For example, Yin et al.(Yin and Afsari, 2023) developed an ontology-aided parser that translates questions into SPARQL queries, though it handled only a narrow set of query types. Other works adopted machine learning pipelines, for example, Nabavi et al.(Nabavi and Becerik-Gerber, 2023) combined SVM-based query classification with keyword extraction, yet such solutions often required extensive manual setup and lacked flexibility. Recent research has shifted toward conversational AI and transformer-based language models. Zheng and Fischer (Zheng and Fischer, 2023) introduced *BIM-GPT*, a dynamic prompt-based virtual assistant framework that integrates BIM databases with GPT for natural-language queries. Their system uses a prompt manager to translate user questions into structured prompts, allowing the LLM to interpret the query, retrieve relevant BIM information, and generate a summary answer. Hellin et al. (Hellin et al., 2025) also presented an LLM-based agent workflow that interacts directly with IFC models. By avoiding heavy ontology pre-processing, their method correctly answered around 80% of queries across architectural, structural, and MEP datasets. These studies demonstrate that LLMs can overcome some rigidity of earlier methods, handling complex multi-hop questions and open-ended queries that previous systems struggled with. Researchers have also begun exploring multi-agent systems (MAS) driven by LLMs to tackle more complex BIM tasks. Du et al.(Du et al., 2024) proposed *Text2BIM*, an LLM-based multi-agent framework that generates 3D building models from natural language descriptions. In their approach, multiple specialized LLM agents collaborate: one agent refines the user’s design prompt, another writes script code for model creation, and another checks the resulting model against design rules, iterating until a consistent BIM model is produced. These multi-agent frameworks illustrate the growing interest in modular conversational agents in AEC, where each agent can specialize (e.g., one for geometry, one for data queries, one for validations) and collectively handle complex, multi-step tasks. However, existing systems tend to focus on specific applications such as model generation, compliance check, or coordination without offering a conversational assistant capable of querying and checking both semantic attributes and geometric data within BIM models. Moreover, many of these approaches rely heavily on precise prompt engineering and often struggle with scalability due to the token limitations inherent in current LLM architectures. When dealing with large-scale BIM models that contain extensive elements and metadata, it becomes impractical to pass the entire context to the LLM, leading to incomplete or inaccurate results. The proposed solution addresses these limitations by introducing a modular multi-agent framework that enables efficient and scalable retrieval from complex BIM models without overloading the LLM, while preserving access to both attribute-level and geometric reasoning capabilities (Kiavarz et al., 2023, Usmani et al., 2021, Hor et al., 2018).

4. Contribution

This research presents a novel conversational multi-agent platform for BIM that bridges several gaps identified in prior works. In contrast to earlier BIM chatbots or question-answering systems that focused solely on attribute queries or textual metadata, our platform integrates both semantic data retrieval and geometric reasoning within a unified chat interface. This allows users to issue natural language queries not only about properties and classifications (e.g., “What is the material of the walls on Level 2?”) but also about spatial conditions and metrics (e.g., “Highlight elements adjacent to the north façade” or “Compute the total volume of concrete in columns”). The proposed method is integrated with the Autodesk Construction Cloud (ACC) for readability and scalability. A key contribution of our work is the modular multi-agent architecture, where each agent specializes in a specific function such as querying, filtering, validation, computation, spatial analysis, or visualization and is coordinated through a central controller. Unlike prior monolithic or prompt-dependent implementations, this design allows scalable, interpretable, and extensible orchestration of BIM-related tasks. New agents can be added incrementally to support additional workflows without disrupting the system as a whole. This structure enhances transparency, reduces failure propagation, and aligns with real-world BIM practices that require separation of concerns. Additionally, our system is distinguished by its ability to operate effectively on large-scale BIM models without being constrained by the token limits of LLMs. Existing LLM-driven solutions often attempt to embed entire BIM contexts or JSON representations within prompts, an approach that breaks down as model size and complexity increase. Furthermore, the platform includes a tightly integrated with ACC 3D viewer that visually responds to queries by isolating, highlighting, or zooming to relevant model elements. This addresses a major limitation of existing approaches, which provide text-based responses but lack visual feedback tied to the BIM.

5. Methodology

The proposed system is built on a modular Multi-Agent System (MAS) architecture that enables natural language interaction with BIM models while supporting semantic querying, geometric computation, and spatial reasoning. The system is designed to overcome the limitations of monolithic LLM-based approaches by distributing tasks across specialized agents coordinated by a central Supervisor Agent.

5.1 System Overview

The architecture consists of three main layers:

- **BIM Integration Layer:** Provides access to geometric and semantic BIM data.
- **Agentic System Layer:** Contains specialized agents responsible for executing domain-specific tasks.
- **Interaction Layer:** Enables user interaction through a conversational interface and 3D visualization.

This layered design improves the scalability, interpretability, and extensibility of the system.

5.1.1 BIM Integration Layer The BIM Integration Layer connects the platform to Autodesk Construction Cloud (ACC) using Autodesk Platform Services (APS), including the Data Management and Model Derivative APIs. The system supports multiple BIM formats, including RVT, DWG, DGN, IFC, and NWD, and enables querying of model elements by element_Id, type, or property. This unified integration ensures consistent access to BIM data regardless of file format or source.

5.1.2 Agentic System Architecture The system follows a modular agent-based design where each agent performs a specific function and is orchestrated by a central Supervisor Agent (Fig. 1). This approach extends prior LLM-based BIM systems by enabling structured task decomposition and tool-based execution (Zheng and Fischer, 2023, Hellin et al., 2025, Du et al., 2024).

The Supervisor Agent receives a user query Q, decomposes it into sub-tasks, and selects the most appropriate agent A_i from the agent set $\mathcal{A}$:

$$A^* = \arg\max_{A_i \in \mathcal{A}} P(A_i \mid Q) \tag{1}$$

where $P(A_i \mid Q)$ represents the likelihood of agent A_i being suitable for the task (Yanwei Yue, 2025).

Each agent operates within a defined scope:

- **Query Agent:** Extracts element attributes based on user input (e.g., “What is the material of walls on level 2?”).
- **Filter Agent:** Filters elements based on constraints (e.g., “Show all elements where volume is more than 1000 m^3”).
- **Validation Agent:** Detects missing or inconsistent attributes (e.g., “Find missing values in Unit of Measurement (UoM)”).
- **Computation Agent:** Computes geometric properties such as volume, area, and length.
- **Summary Agent:** Aggregates statistics (e.g., “How many walls exist per floor?”).
- **Spatial Agent:** Handles spatial reasoning tasks such as adjacency, intersection, and containment.
- **Highlight Agent:** Sends visualization commands to the 3D viewer to isolate, color, or zoom into elements.

This modular structure enables extensibility and reduces error propagation across tasks.

5.1.3 Interaction Layer A user-friendly conversational interface is integrated with a 3D viewer to enable intuitive interaction with BIM data. The system returns both textual responses and visual feedback, where queried elements are highlighted, isolated, or zoomed within the model.

This multimodal interaction improves interpretability and allows users to validate results directly within the BIM environment.

https://doi.org/10.5194/isprs-archives-XLIX-B4-2026-501-2026 | © Author(s) 2026. CC BY 4.0 License.

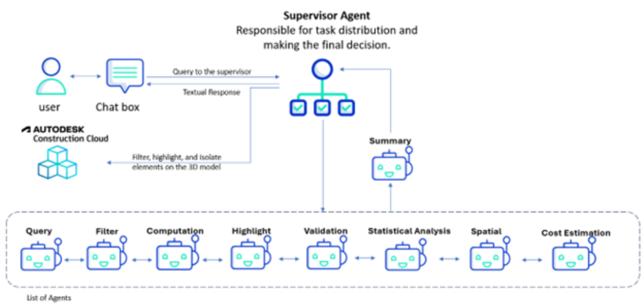


Figure 1. System architecture of the proposed BIM multi-agent platform.

5.2 Geometric Computation

The Computation Agent extracts physical properties from BIM geometry. For standard elements, Boundary Representation (B-Rep) methods are used. For complex geometries, a voxel-based approximation is applied.

A voxel V is uniquely defined as:

$$V_{ID} = X_{ID} + Y_{ID} \cdot DimX + Z_{ID} \cdot DimX \cdot DimY \quad (2)$$

The total volume is computed as:

$$V_{total} = \sum_{i=1}^{N} o_i \cdot V_{voxel} \quad (3)$$

where o_i represents voxel occupancy.

This approach allows for robust computation for irregular geometries and supports advanced spatial analysis.

5.3 Spatial Reasoning

The Spatial Agent evaluates relationships between BIM elements using geometric and topological rules. Key spatial operations include:

- **Intersection:** $ST\_Intersects(A, B)$
- **Containment:** $ST\_Contains(A, B)$
- **Adjacency:** Defined using a distance threshold ϵ
- **Touching:** $ST\_Touches(A, B)$

Adjacency is defined as:

$$d(S_1, S_2) < \epsilon \quad (4)$$

where d is the minimum distance between elements.

These relationships are derived from svf2 format from ACC.

6. Evaluation

Evaluating conversational BIM systems requires assessing both natural language understanding and the correctness of domain-specific computations. Traditional LLM benchmarks are insufficient, as BIM tasks involve structured data retrieval, geometric reasoning, and spatial analysis.

6.1 Evaluation Method

The proposed system is evaluated using a task-based framework inspired by BIM-specific benchmarks such as IFC-Bench (Hellin, 2025). The evaluation focuses on four categories of increasing complexity:

- **Level 1, Direct Retrieval:** Simple attribute extraction (e.g., material, element type).
- **Level 2, Filtered Queries:** Conditional selection of elements based on attributes.
- **Level 3, Aggregation and Computation:** Quantitative analysis such as counts, volumes, or areas.
- **Level 4, Spatial Reasoning:** Queries involving adjacency, containment, or geometric relationships.

Each query is evaluated against a ground truth derived from the BIM model using deterministic methods except the semantic search part which is non-deterministic.

6.2 Evaluation Metrics

To assess system performance, the following metrics are defined:

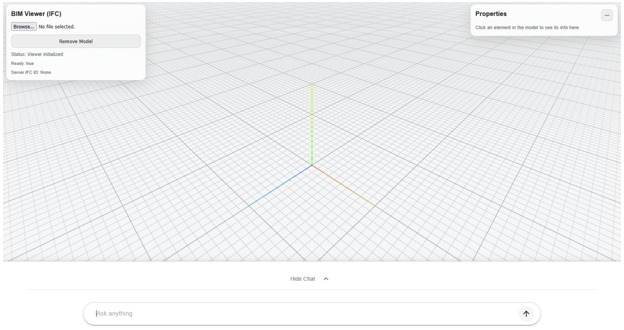
- **Numerical Accuracy:** For quantitative output, an error metric is computed as:

  $$\text{Error (\%)} = \frac{|V_{pred} - V_{true}|}{V_{true}} \times 100 \quad (5)$$

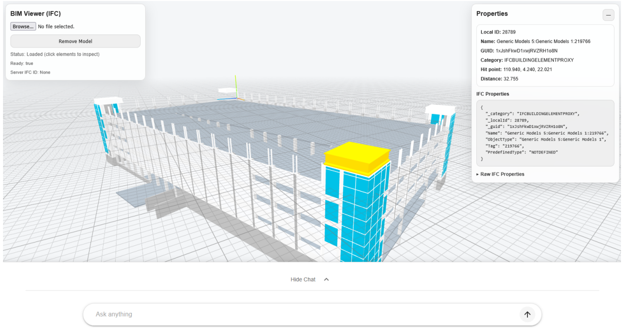
  Results are considered:

  - **Correct** if error $< 5\%$
  - **Partially Correct** if error is between 5% and 20%
  - **Incorrect** if error $> 20\%$

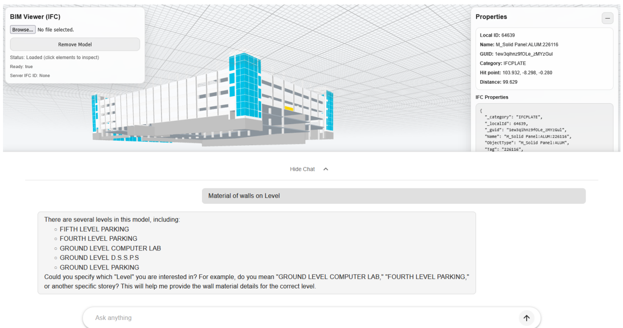
- **Semantic Equivalence:** Textual responses are considered correct if they convey the same meaning as the ground truth.



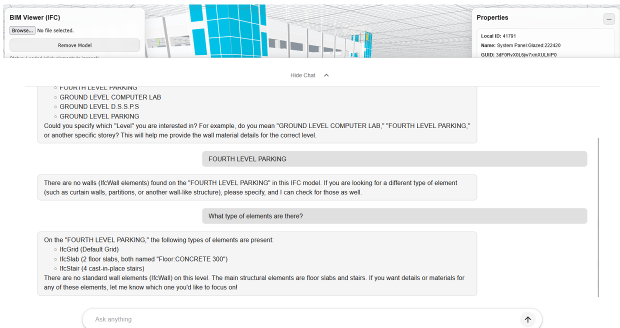
(a) Overview of the proposed platform interface. The interface includes a 3D model viewer and a chat-based interaction panel. Users can load a model via an ACC link or upload an IFC file from the upper-left panel.



(b) Model interaction and property inspection after loading. Users can navigate the model and inspect semantic attributes assigned to elements.



(c) Example of conversational clarification. The system asks for missing information, demonstrating session memory and multi-turn interaction.



(d) Example of semantic querying. The system identifies IFC element types present on a selected level and summarizes the results.

Figure 2. Examples of the proposed conversational BIM platform interface and interaction workflow.

- **Tool Execution Validity:** Whether the correct agent/tool was used to produce the result.

Each query is labeled as: Correct, Partially Correct, or Incorrect.

7. Results and Discussion

Table 1 presents a representative set of evaluation queries across different complexity levels.

The results demonstrate that the proposed multi-agent system achieves high accuracy across different query types, particularly for direct retrieval and filtered queries (Levels 1 and 2), where deterministic data access ensures near-perfect performance.

For aggregation and computation tasks (Level 3), minor deviations are observed due to geometric approximation methods such as voxelization. However, these deviations remain within acceptable tolerance thresholds, indicating reliable performance for practical applications.

Spatial reasoning tasks (Level 4) present greater challenges, as they depend on geometric precision and threshold definitions (e.g., adjacency tolerance ϵ). Slight discrepancies in these cases highlight the sensitivity of spatial computations to model quality and geometric representation.

The modular MAS architecture plays a critical role in maintaining performance by isolating tasks across specialized agents. Compared to monolithic LLM approaches, this reduces context overload and improves interpretability of results. The Supervisor Agent ensures effective task decomposition and routing, minimizing reasoning errors.

Another key advantage is the integration of visual feedback through the 3D viewer. This allows users to verify results directly within the BIM environment, addressing a major limitation of text-only systems.

8. Limitations

Despite the promising results, several limitations remain in the proposed framework. The system relies heavily on the completeness and consistency of BIM data. Missing attributes, inconsistent naming conventions, or poorly structured models can directly affect the accuracy of results. Geometric approximation errors such as voxel-based or simplified geometric representations may introduce minor inaccuracies in computed quantities such as volume or area. Moreover, Spatial reasoning tasks depend on threshold parameters (e.g., adjacency tolerance ϵ). Small variations in these parameters or model precision can lead to different results. Also, the orchestration of multiple agents introduces additional computational overhead compared to single-step query systems, particularly for complex queries requiring sequential reasoning. Although raw BIM data is not exposed, the system depends on LLMs for intent parsing. Ambiguous or poorly structured user queries may lead to incorrect task decomposition.

9. Future Work

Future research will focus on enhancing the robustness, scalability, and intelligence of the proposed system. Incorporating

Table 1. Evaluation results of the proposed BIM multi-agent system

ID	Query	Ground Truth	Model Output	Result
1	Material of walls on Level 2	Concrete	Concrete	Correct
2	Columns with fire rating $>$ 2 hrs	12	12	Correct
3	Number of walls on Level 1	45	45	Correct
4	Elements with volume $>$ 1000 m^3	8	8	Correct
5	Total volume of columns	1250 m^3	1210 m^3	Correct
6	Missing UoM attributes	6	6	Correct
7	Slab area on Level 3	980 m^2	950 m^2	Correct
8	Beam types	4	4	Correct
9	Doors per floor	Exact counts	Minor mismatch	Partial
10	Walls adjacent to north facade	15	14	Partial
11	Elements intersecting HVAC	9	9	Correct
12	Rooms connected to corridor	6	6	Correct
13	Total slab volume	2300 m^3	2280 m^3	Correct
14	Elements in Zone A	22	22	Correct
15	Structural materials	Steel, Concrete	Steel, Concrete	Correct
16	South-facing windows	18	17	Partial
17	Elements touching roof slab	11	11	Correct
18	Empty attributes count	35	32	Partial
19	Avg wall thickness	0.30 m	0.31 m	Correct
20	Elements in Level 2	Correct set	Correct set	Correct
21	Total beam length	540 m	530 m	Correct
22	Floors with fire-rated walls	3	3	Correct
23	Columns per structural zone	Exact counts	Slight mismatch	Partial
24	Elements intersecting slab	14	13	Partial
25	Total wall area	3200 m^2	3150 m^2	Correct
26	Mechanical equipment count	27	27	Correct
27	Rooms with no windows	5	5	Correct
28	Elements within 5m of core	19	18	Partial
29	Max column height	12 m	12.2 m	Correct
30	Slabs touching exterior walls	10	9	Partial

more advanced geometric kernels and precise spatial indexing methods can improve the accuracy of spatial queries and reduce sensitivity to threshold parameters. Developing learning-based routing mechanisms can optimize agent selection and reduce latency by dynamically adapting to query patterns. Fine-tuning LLMs on BIM-specific datasets can improve query understanding and reduce ambiguity in user intent. Establishing standardized BIM evaluation datasets and benchmarks will allow for more rigorous comparison across different AI-driven BIM systems.

10. Conclusion

This paper presented a novel MAS framework for enabling natural language interaction with BIM data. By integrating LLMs with specialized agents, the proposed system supports semantic querying, geometric computation, and spatial reasoning within a unified and interpretable architecture. Unlike monolithic LLM-based approaches, the proposed framework decomposes complex user queries into structured sub-tasks executed by dedicated agents. This modular design improves scalability, reduces reasoning errors, and enables transparent interaction with BIM models. The integration with ACC and support for different BIM formats further enhances the applicability of the system in real-world workflows. The evaluation results demonstrate that the system achieves high accuracy across a range of query types, including attribute retrieval, filtering, aggregation, and spatial analysis. Although minor deviations were observed in geometry-based computations and spatial queries, the results remain within acceptable tolerance levels for practical use. The proposed approach contributes to bridging the gap between human-centric natural language interaction and structured BIM data systems. By allowing intuitive access to complex building data, the framework has the potential to improve productivity, reduce manual effort, and support data-driven decision-making. Overall, this work establishes a foundation for intelligent, agent-driven BIM systems and highlights the potential of combining LLMs with domain-specific multi-agent architectures. Future advances in spatial reasoning, agent coordination, and domain-adapted language models are expected to further enhance the capabilities of such systems.

References

Du, Y., Zhang, Z., Gao, X., 2024. Text2BIM: A multi-agent framework for generating BIM models from design intent in natural language. *Journal of Computing in Civil Engineering*, 38(2), 04023085.

Hellin, M., Karakhan, A., Ramos, F., 2025. IFC-LLM: Interpretable large language model agents for open BIM query workflows. *Journal of Information Technology in Construction (ITcon)*, 30, 45–61.

Hellin, S., 2025. Ifc-bench: Benchmark dataset for bim ai systems.

Hor, A.-H., Sohn, G., Claudio, P., Jadidi, M., Afnan, A., 2018. A semantic graph database for BIM-GIS integrated information model for an intelligent urban mobility web application. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 4, 89–96.

Kiavarz, H., Jadidi, M., Esmaili, P., 2023. A graph-based explanatory model for room-based energy efficiency analysis based on BIM data. *Frontiers in Built Environment*, 9, 1256921.

Nabavi, S., Becerik-Gerber, B., 2023. A machine learning pipeline for natural language interaction with BIM. *Advanced Engineering Informatics*, 58, 101283.

Usmani, A., Jadidi, M., Sohn, G., 2021. Towards the automatic ontology generation and alignment of BIM and GIS data formats. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 183–188.

Yanwei Yue, Guibin Zhang, B. L., 2025. MasRouter: Learning to Route LLMs for Multi-Agent Systems. *https://arxiv.org/abs/2502.11133*.

Yin, R., Afsari, K., 2023. Ontology-aided semantic parsing for natural language queries over BIM data. *Automation in Construction*, 149, 104766.

Zheng, J., Fischer, M., 2023. BIM-GPT: A virtual assistant for natural language interaction with BIM using prompt engineering. *Automation in Construction*, 154, 104961.